

Sabine Imeri, Wjatscheslaw Sterzer, Matthias Harbeck

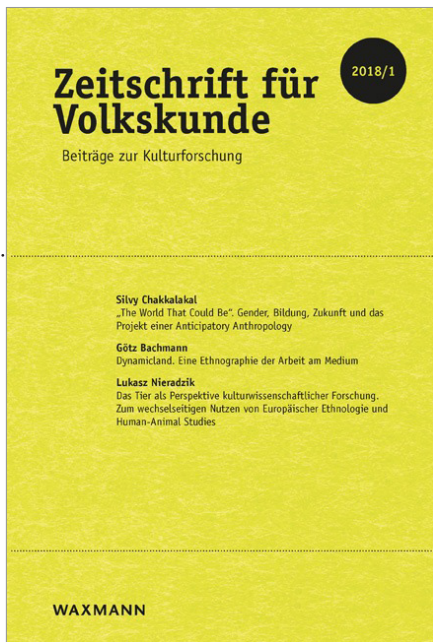
Forschungsdatenmanagement in den ethnologischen Fächern. Bericht aus dem Fachinformationsdienst Sozial- und Kulturanthropologie

Aus:

Zeitschrift für Volkskunde

114. Jahrgang, Heft 1/2018

ISSN 0044-3700



www.waxmann.com/zfv



WAXMANN

Steinfurter Straße 555
48159 Münster

Tel.: 0251 / 265 04-0

Fax: 0251 / 265 04-26

www.waxmann.com

info@waxmann.com

facebook.com/WaxmannVerlag

© Waxmann Verlag GmbH, 2018

Forschungsdatenmanagement in den ethnologischen Fächern. Bericht aus dem Fachinformationsdienst Sozial- und Kulturanthropologie

Nach dem Willen von nationalen wie internationalen Forschungsförderern sollen Daten aus öffentlich geförderter Forschung künftig systematisch langfristig archiviert, möglichst frei ausgetauscht und zur weiteren Verwendung zur Verfügung gestellt werden. Neben der dauerhaften Sicherung soll so zum einen das Versprechen nach mehr Transparenz und Teilhabe eingelöst werden, zum anderen und vor allem aber sind die Förderinstitutionen an mehr Effizienz und Überprüfbarkeit interessiert; sie haben entsprechend längst begonnen, in ihre Richtlinien die Forderung nach zumindest temporärer Archivierung und Möglichkeiten der sekundären Nutzung von Daten

einzuarbeiten. Künftig wird also bei der Antragstellung vermehrt die Vorlage sog. Datenmanagementpläne verlangt werden; Antragstellende werden sich zum Verbleib der Daten nach dem Ende der Forschung äußern müssen. Während wissenschaftspolitisch aktive Gremien und datenintensive Wissenschaften die Etablierung von Regeln und Standards des sogenannten Forschungsdatenmanagements (FDM) vorantreiben, technische Infrastrukturen aufgebaut und Workflows entwickelt werden, hat sich dieser Diskurs in den ethnologischen Fächern bislang kaum niedergeschlagen. Zum Beispiel verfügen die Ethnologien in Deutschland bisher nicht über fachspezifische, von den Fachgesellschaften getragene Positionspapiere und Empfehlungen zum Umgang mit Forschungsdaten. Überdies sind konkrete Anwendungskenntnisse und vertieftes Wissen wenig verbreitet – eine erste Erkenntnis aus den Arbeiten des Fachinformationsdienstes (FID) Sozial- und Kulturanthropologie an der UB der Humboldt-Universität, der sich seit Juni 2016 mit Fragen des FDM für die ethnologischen Fächer beschäftigt.¹ Der FID setzt sich zunächst mit der gängigen Praxis im Umgang mit Forschungsdaten auseinander und erarbeitet Schwierigkeiten und Anforderungen, die aus den Besonderheiten der ethnografischen Forschung erwachsen. Wie könnte ideales Forschungsdatenmanagement für die ethnografische Forschung künftig aussehen? Welche Bedingungen müssten erfüllt sein? Und wie ließe sich das in Forschungsroutinen integrieren?

Der FID bearbeitet solche und ähnliche Fragen in Zusammenarbeit mit Forschungsprojekten und Fachgesellschaften, mittels Umfragen, Interviews und Gesprächen. Zudem gelten Recherchen den technischen wie organisatorischen Lösungen, die etablierte Datenarchive für ähnliche Probleme gefunden haben. Denn während GESIS – Leibniz-Institut für Sozialwissenschaften bereits seit mehreren Jahrzehnten Daten aus quantitativer Sozialforschung archiviert und zur Sekundärnutzung bereitstellt, stehen für Daten aus ethnografischer oder allgemeiner aus qualitativer (Sozial-)Forschung in Deutschland kaum Möglichkeiten der dauerhaften Archivierung zur Verfügung. Der FID zielt in der gegenwärtigen ersten Förderphase nicht auf die Etablierung eines eigenen Datenarchivs, vielmehr sollen infrastrukturelle Anforderungen aus vorhandenen Workflows und Empfehlungen abgeleitet werden, um zur Konzeptualisierung fachspezifischer Lösungsansätze beizutragen. Dieser Bericht gibt Einblicke in erste Zwischenergebnisse aus diesen Arbeiten, auch wenn hier zunächst mehr Fragen aufgeworfen als konkrete Antworten gegeben werden.

Es hat sich aus unserer Sicht bewährt, drei Ebenen des Forschungsdatenmanagements zu unterscheiden und deren Erfordernisse getrennt zu betrachten (auch wenn sie zu einem Prozess gehören): *Erstens* das prozessbegleitende Datenmanagement – also etwa Ordnung und Ablage von Daten, projektinterner Austausch in Gruppen oder Datensicherung –, das bereits technische oder rechtliche Fragen und

1 Der Anstoß dazu kam aus der Fachcommunity heraus, die in die Antragstellung – der FID wird projektförmig von der DFG gefördert – involviert war. Bereits hier waren einerseits wissenschaftspolitische Notwendigkeiten, andererseits aber auch die Nachnutzungspotenziale im Blick.

Probleme aufwerfen kann, z.B. beim Datenschutz. Außerdem fehlt es vielen Institutionen an technischen Voraussetzungen adäquater Umsetzung, wenn z.B. in internationalen Kooperationsprojekten Daten grenzüberschreitend gemeinsam genutzt werden sollen. *Zweitens* die Langzeitarchivierung, die klären muss, wie und wo, unter welchen technischen Bedingungen und rechtlichen Vorgaben Daten nach dem Ende eines öffentlich geförderten Projektes dauerhaft abgelegt werden können. Die DFG hat bereits 1998 eine Aufbewahrungsfrist von mindestens zehn Jahren festgelegt, ansonsten ist offen, über welche Zeiträume hier gesprochen werden kann oder sollte. Langzeitarchivierung meint jedenfalls nicht die Speicherung auf externen Festplatten oder bei (kommerziellen) Cloud-Anbietern, sondern mindestens auf institutionellen Servern oder eben bei professionellen Datenarchiven, die u.a. eine „Lesbarkeit“ der Daten mit zukünftigen Softwaregenerationen gewähren sollen. Dass Daten dort auch gefunden werden können, führt *drittens* zur Nachnutzung, die Dritten die Verwendung von Daten für neue Forschungen ermöglichen soll. Solche Nachnutzungen sind für die ethnografische Forschung schwieriges Terrain: Schließlich gelten die Anforderungen verantwortungsvollen Forschens auch für den Umgang mit Daten und deren Verwendung nach dem Abschluss einer Forschung.

Insgesamt stehen Ethnologinnen und Ethnologen dem Datenmanagement zwar häufig kritisch, aber nicht grundsätzlich ablehnend gegenüber. Das Spektrum der Einschätzungen vor allem zu Langzeitarchivierung und Nachnutzung von Daten durch Dritte ist gleichwohl enorm breit: Stark ablehnende Haltungen begegnen hier ebenso wie Nachfragen nach konkreten infrastrukturellen Angeboten zur Datenarchivierung. Einwände haben zunächst eine forschungsökonomische Dimension, beziehen sich also auf den zu erwartenden Arbeitsaufwand oder die Finanzierung der Datenaufbereitung. Auch wenn bisher noch nicht erkennbar ist, wie sich neue Routinen des Datenmanagements in Forschungspraxis werden übersetzen lassen: Weil sie den gesamten Forschungsprozess sowie forschungsethische Fragen tangieren, scheint es, als wären das keine Arbeiten, die sich ohne weiteres delegieren ließen. Im Folgenden sollen drei der vorgefundenen Problemlagen kurz ausgeführt werden.

1. Erhebungskontexte dokumentieren: Ob und wie Feldforschungsbeziehungen und die komplexen Kontexte der Datenerhebung so dokumentiert werden können, dass sie für Dritte – vielleicht sogar fachfremde Forschende – möglichst transparent, verständlich und damit überhaupt sinnvoll nutzbar werden, ist eine offene Frage: „Datasets don't speak for themselves“ [Rena Lederman], zumal in den offenen, kaum standardisierten Forschungsprozessen der Ethnologien die Heterogenität sich oft gegenseitig kommentierender Forschungsdaten erheblich ist. Diskutiert werden müssten entsprechend fachliche Standards für die angemessene Dokumentation von Erhebungskontexten. Die Anreicherung mit Metadaten allein wird hier kaum genügen, überdies fehlen derzeit spezifische Metadatenstandards und einheitliche Fachthesauri.

2. Daten für die Archivierung auswählen: Die Bereitschaft von Forschenden, Daten in ein Datenarchiv zu überführen, nimmt ab, je mehr auch die forschende Person im Material erkennbar wird. Selbstschutz und Vorbehalte gegenüber Überprüfung und Kontrolle sind häufige Bezugspunkte. Besonders mit Blick auf hybrides Material, wie Feldtagebücher, ist die Skepsis groß. Genereller geht es aber um die Auswahl von zur Nachnutzung geeigneten Daten und deren Aufbereitung mit vertretbarem Aufwand. Hierfür müssen fachspezifische Kriterien erst noch entwickelt werden. Feldtagebücher werden möglicherweise nicht oder nur unter sehr spezifischen Bedingungen dazu gehören.

3. Rechtliche Fragen und Forschungsethik: Das mit Blick auf Langzeitarchivierung und Nachnutzung wichtigste und kontroverseste Thema ist die Wahrung der Vertraulichkeit in Verbindung mit Fragen von Daten- und Persönlichkeitsschutz und Forschungsethik. Strittig ist etwa, ob und wie in offenen Feldsituationen Einwilligungserklärungen eingeholt werden können, die dann auch auf Szenarien zukünftiger Nutzung der Daten durch unbekannte Dritte ausgedehnt werden müssen. Lösungen im Spannungsfeld von Anonymisierung komplexen ethnografischen Materials und dem Erhalt von dessen Interpretierbarkeit wird hier eine zentrale Rolle zukommen. Zumal sich Konzepte von „Privatheit“ oder „Sensibilität“ dynamisch entwickeln und bei Datenübergabe an ein Archiv kaum abzusehen ist, auf welchen Wegen unvollständig anonymisierte Daten künftig Schaden anrichten können. Bestenfalls andiskutiert sind Fragen des Urheberrechts, auch mit Blick auf kollaborativ erzeugte Daten und Cultural Property.

Trotz vieler Unklarheiten sind viele Forschende bereit, Daten künftig *unter bestimmten Bedingungen* in Repositorien zugänglich zu machen. Abschließend sollen deshalb Anforderungen an künftige Datenarchive in wenigen Strichen skizziert werden. Weil die „Offenheit“ ethnografischer Daten Grenzen hat, Datenpublikationen im eigentlichen Sinn nur in wenigen Fällen möglich sein werden, sind die Anforderungen an die Vertrauenswürdigkeit von Datenarchiven hoch und vielfältig: Benötigt wird etwa ein umfassendes Rechtemanagement, das nicht intendierte Verwertungen von Daten verhindert, gestufte Zugangsrechte und vor allem Zugangskontrolle zuverlässig gewährleistet. Datenarchive, die Daten aus qualitativer Forschung beherbergen – wie z.B. das Datenarchiv Qualiservice an der Universität Bremen –, arbeiten bereits mit gestaffelten Zugriffsrechten. In Abhängigkeit von Datentyp und Anonymisierungsgrad sind dort Schutzstufen bis hin zu sog. *Safe Rooms* – in denen hochsensible Daten unter kontrollierten Bedingungen nur an Gastarbeitsplätzen vor Ort zur Verfügung gestellt werden – etabliert worden. Auch regeln Nutzungsverträge Rechte und Pflichten. Es gibt also bereits erprobte Verfahren, die künftig für Daten aus ethnografischen Forschungen Anwendung finden könnten. Viele Forschende möchten zudem selbst entscheiden können, ob, an wen und wofür ihre Daten zur Nachnutzung

freigegeben werden. Auch aus der Perspektive der Nachnutzung kann die Möglichkeit, Kontakt zu den Primärforschenden aufzunehmen, wichtig sein. Ebenso sollten Exit-Strategien bzw. Möglichkeiten diskutiert werden, archivierte Daten wieder zurückzuziehen, etwa wenn sich die Bedingungen ändern, unter denen die Daten übergeben wurden. Insgesamt scheinen generische Lösungen für die ethnologischen Fächer unzureichend zu sein. Datenarchive sollten vielmehr einen möglichst hohen Grad an disziplinspezifischer Professionalisierung aufweisen.

In Anbetracht der wissenschafts- und förderpolitischen Brisanz bei relativem Mangel an Wissen in den ethnologischen Fachcommunitys bieten die Arbeiten des FID die Möglichkeit, das Thema vermehrt auf die Agenda zu bringen und zu diskutieren. Angestoßen ist damit ein Prozess, an dessen Ende zum einen eine differenzierte, den Eigenheiten der Ethnologien gerecht werdende Positionierung zu den Forderungen der Forschungsförderer und zum anderen adäquate Anforderungen an Datenrepositorien und Empfehlungen für die Umsetzung in Forschungspraxis stehen können.

Sabine Imeri, Wjatcheslaw Sterzer, Matthias Harbeck